Tibetan Language Kit for Macintosh and Its Use at the Otani University Shin Buddhist Comprehensive Research Institute

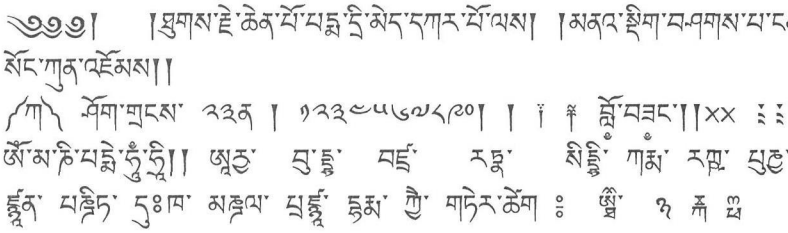
Tibetan Works Research Project
Otani University Shin Buddhist Comprehensive Research Institute

The use of the personal computer in Tibetan studies have recently become quite popular. Many scholars are making documents and databases in Tibetan, frequently using Tibetan letters. There are several Tibetan script systems now available. This paper introduces the Tibetan Language Kit (hereafter abbreviated as TLK), a new Tibetan script for the Macintosh personal computer developed by the Otani University Shin Buddhist Comprehensive Research Institute, and its use at the institute.

Outline of the TLK

An earlier version of the TLK for Macintosh called the Druk system was developed in the late 1980's at the request of, and with the funding of, the Royal Government of Bhutan. With the evolution of the operating system for Macintosh from Version 6 to Version 7.1 in the early 1990's, it became necessary to upgrade the Druk system. The project of upgrading the system was taken on by the Shin Buddhist Comprehensive Research Institute, in collaboration with Dr. Yoshiro Imaeda (CNRS, France), Mr. Steven Hartwell (Multiscript Solutions International, Paris, France) and Mr. Yoichi Fukuda (Toyo Bunko, Tokyo, Japan).

The TLK is compatible with the Macintosh Worldscript multi-script computing environment, allowing the user to type and print Tibetan script and text with other scripts (Roman, Japanese, Chinese, etc.) in the same document, and constitute databases in Tibetan. TLK contains an outline font of high quality called Kailasa which can display and print at various sizes all Choekey letters and combinations as well as letters from the most common Sanskrit and Terma texts.



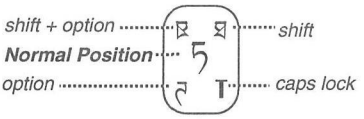
Input Methods of TLK

TLK for Macintosh has an easy and consistent method for Tibetan text entry which allows one to type Tibetan letters and signs directly. With this method, one just types the independent radical characters in a certain sequence, leaving to system to automatically display the complex vertical stacks (combinations). The system also handles automatic word wrap and selection of syllables, a feature which facilitates greatly the layout of the text.

Keyboard Layouts

TLK has three keyboard layouts: TibetanUS, OtaniUS and TibWriter.

(1) TibetanUS



Consonants: As shown in the above chart, the majority of Tibetan characters are allocated to the keys of their Roman equivalent in the Wylie transcription system. Tibetan signs and certain characters which do not have appropriate equivalents in Roman keyboard are allocated to somewhat arbitrary keys. The shift and option keys allow one to type characters and signs which are not allocated in the keyboard in its normal position. In principle, the shift key allows one to type the aspirated counterparts of the non-aspirated consonants. Likewise, using the option key, one can type the reversed counterparts of the consonants. The combination of the shift and option keys is used to type the aspirated reversed consonants.

Tibetan Conjunct Character: There are many conjunct characters in Tibetan. The conjunct character is a special character used for constructing two or more consonants into a vertical stack. The conjunct character is automatically constructed by typing the conjunct key between the typing of

each consonant. In order to type སྒྲ, for example, we first type ས, then the conjunct key, next ཀ, then the conjunct key again, and finally ར.

Vowels: The four vowel signs *gigu* (ི), *shapkyu* (ུ), *drengo* (ེ) and *naro* (ོ) are allocated to the keys of their Roman vowel equivalents. The reversed *gigu* is obtained by typing option+*gigu*. Double vowel signs are obtained by typing twice their respective single vowel signs.

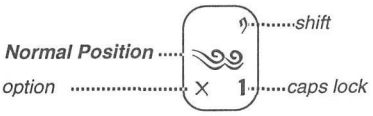
Numbers: The Tibetan numbers are located on the top row.

Punctuation marks: *Tsek*(་) is on the spacebar, and *shey* (།) is next to the shift key on the right side of the keyboard.

(2) OtaniUS

The OtaniUS keybord is nearly identical to the TibetanUS keybord. Those who wish to input through the Wylie transcription system will find it convenient to use the OtaniUS keyboard. No Tibetan character is allocated in the shift position. Therefore, except the reversed letters (which are obtained by the option key in the same way with the TibetanUS keybord), all Tibetan consonants can be otained in the normal position.

(3) TibWriter



The allocation of Tibetan characters follows as much as possible the keyboard of a Tibetan mechanical typewriter. Each of the Tibetan characters are allocated to a key in the normal position as shown in the chart above. No shift or option keys are required to input Tibetan characters using this keyboard.

Entering the Tibetan Text

Tibetan syllables can include several elements which superimpose on the base character (=radical), such as vowel signs and super- or sub-joined characters.

There are circumstances in which it is desirable to use a unique input sequence, such as when one eventually wishes to use the global search/replace in a word processor, or in creating a database where sorting is important. If syllables were entered in different orders (for example, sometimes *ka* then *gigu* and at other times *gigu* then *ka*), then by searching for the syllable *ki*, one would not find the syllable typed *gigu* then *ka*, even though they would appear the same on the screen. Also, such syllables might sort differently in the database.

To avoid this dilemma, TLK supports only one ordering of characters in a vertical stack:

(i) a consonant,
(ii) one or more sequences of the conjunct character and a consonant (if any),
(iii) a vowel or certain double vowels (if any), and
(iv) one of certain modifications such as *anusvara* (if any).

TLK system cannot actually prevent out of sequence entry, but it can bring attention to it by displaying the syllable with a large black dot which represents the character out of sequence. For example, if the incorrect sequence: *a-anusvara-naro* had been entered, TLK would display the *naro* as a large black dot (like this:ཨཾ•) instead of ཨཾ. (To correct this display, follow the correct sequence: *a-naro-anusvara*.)

Plans for Future Development

The current system does not sort Tibetan syllables in the Tibetan alphabetical order. As the system allows one to incorporate both Choekey and non-Choekey syllables in the same document or database, it is extremely complex to sort these two groups of syllables in a single sequence. But preliminary research and work for making the sorting program is already under way and we hope to be able to provide a sorting program in the future.

At present, the system can display stacks composed of three or four elements at most. In case there are more elements, the system splits them into two (or eventually more) stacks. In this case, the *halant* sign (◌྄ conjunct character) is shown at the bottom of the left stack, for example སྒྲ྄ས སྐྱ྄ན etc. . This is to remind one that although the current system is unable to display and print them in a single stack, the elements splited into

two consecutive stacks are in a single stack in the original document. Within the limits of the available of font space, we hope to enable the system to include more stacks.

The TLK can be obtained free of charge from the Otani University Shin Buddhist Comprehensive Research Institute.

Introduction to the Catalogue Stacks

Otani University possesses many Tibetan works, such as the Peking and Narthang editions of the Kanjur and Tenjur. The Otani University Library has published several catalogues of its Tibetan holdings, including the following:

(1) *A Comparative Analytical Catalogue of the Kanjur Division of the Tibetan Tripitaka* (1932).
(2) *The Tibetan Tripitaka Peking Edition kept in the Otani University, Kyoto: Catalogue and Index* (1960).
(3) *A Comparative Catalogue of the Tanjur Division of the Tibetan Tripitaka* (1965-) volumes 1-7.
(4) *Catalogue of Tibetan Works Kept in Otani University Library* (1973)
(5) *Index to the Catalogue of Tibetan Works Kept in Otani University Library* (1985)

Although the publication of some of these volumes have not yet been completed, the catalogues already published amount to over five thousand pages. If they were to be collected together in one electronic database, they would take up one entire CD disk.

During the course of developing the TLK, we have compiled two Tibetan language electronic catalogues: the Otani Tibetan Collection and the Otani gSer bris. The former is an electronic version of (4) above, while the latter is a newly compiled catalogue. These catalogues were compiled with the help of Mr. Steven Hartwell using Hypercard, the most popular Macintosh database software. As its name shows, Hypercard is a notecard style database. In Hypercard, data inputted into the cards is stored in a file named "Stack."

In creating our electronic catalogues, we have created two stacks: one for inputting information and one for browsing. All data was inputted using the input stack. This input stack has the ability to save data as text. This allows one to save data and also to read all data into the Browser stack. As its name implies, the Browser stack allows one to browse, search, print, mark, sort, etc. (but not input), the data in the stack. The size of the card in the Browser Stack was kept small, facilitating its use on both large and small monitors. Like a card in a library card catalogue, data for each text is contained on one card. Each card is divided into several fields and the data is

contained in the respective fields.

Otani Tibetan Collection Stack

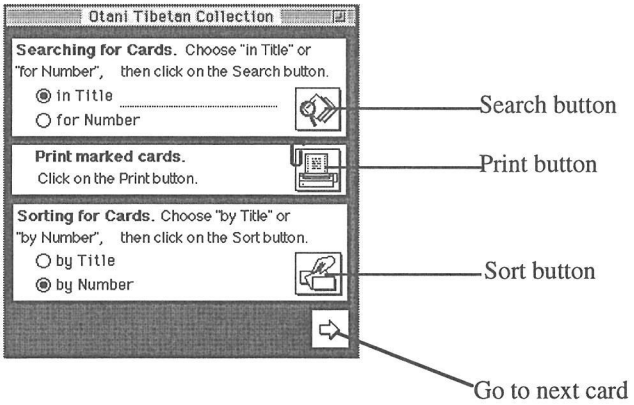
This catalogue stack is a catalogue of Tibetan works in the Otani University Library. The work of inputting the titles of the texts, authors and catalogue numbers has already been completed. However we have yet to create a Browser Stack containing the field for authors. For this reason, the examples of the cards given below will show only two fields: number and title of text.

Outline of the Otani Tibetan Collection Stack

The cards within the Otani Otani Tibetan Collection Stack is arranged in the following order: control card, contents card and cards containing text.

Control Card

The control card is the first card which appears when the stack is opened.



Search, print and sort of data within the stack is done using this card.

Search: Search for data is done in the following manner.

(1) Click either "in Title" or "for Number," then the field appears in the underlined area.

(2) Write the word to search in the underlined area, and click on the Search button. When a corresponding term is found, the card containing that term appears. If there are other cards containg this term, the Search Next button (magnifying glass) on the card shown will be high-

lighted. When one clicks this button, other cards containing this term appear. If there are no corresponding terms, a dialogue box will appear to report this fact.

Print: It is possible to print the data on the marked cards using the Report format supplied with the stack. Print is done in the following way.

(1) Go to the card one want to print.
(2) Click on the Mark button (clip-shaped) on the card. This will mark the card for printing. (When marked, the Mark button on the card will turn color from grey to black.)
(3) When the card is marked, go back to the Control card. Then, click the Print button to begin printing.

Sort: Unfortunately at present the Tibetan letters themselves do not have sortkeys. Thus sortkeys are attached to each card, and it is possible to sort by title using them. Sorting is done in the following manner.

(1) Click either "by Title" or "by Number".
(2) Click the Sort button.

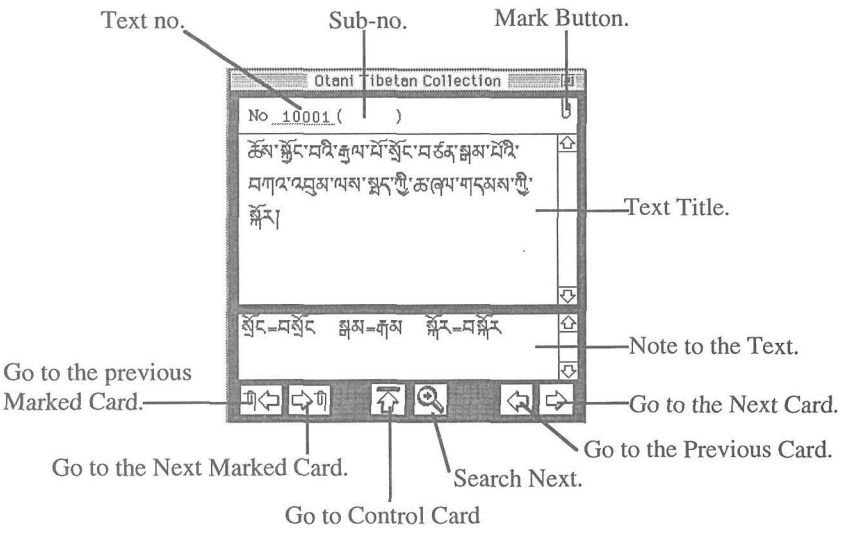
Content Card

When the "go to next card" button on the control card is clicked, the contents card appears.

Thirty letters of the Tibetan alphabet appear on the card. Each letter serves the function of a card box in a library card catalogue system. By clicking a particular Tibetan letter, one can access cards whose title begins with that letter. (When the cards have been sorted by title, the cards appear in alphabetical order.)

Cards Containing Text

Data for each text is contained within a single data card. The following is the card containing data for Otani no. 10001.



Otani gSer Bris Stack

This is a catalogue of the Golden Mansucript Tenjur in Ganden Monastery (abbreviated hereafter as GMTG) which was reproduced photographically by the China National Library and published under the title *Dan-zhu-er* (丹殊爾) from Tian-jin gu-ji chu-ban-she (天津古籍出版社) in 1988. The Otani gSer Bris stack is based on this photographic edition. The manuscript is in 225 *poties*. There is a total of 3972 texts in the Tenjur. The photographic reprint edition is in 100 bound volumes. In this edition, there are 2 or 3 *poties* in each volume.

Outline of the Otani gSer Bris Stack

Just as in the case of the Otani Tibetan Collection Stack, the Otani gSer Bris Stack is composed of (and appear in the order of) the Control Card, Contents Card and cards containing the text. The functions are almost identical to that of the Otani Tibetan Collection Stack.

Control Card

Identical to that of the Otani Tibetan Collection Stack. However, at present, it is impossible to sort by title.

Contents Card

This card lists the seventeen sections of the Tenjur, from *bsTod tshogs* to *dKar chag*.

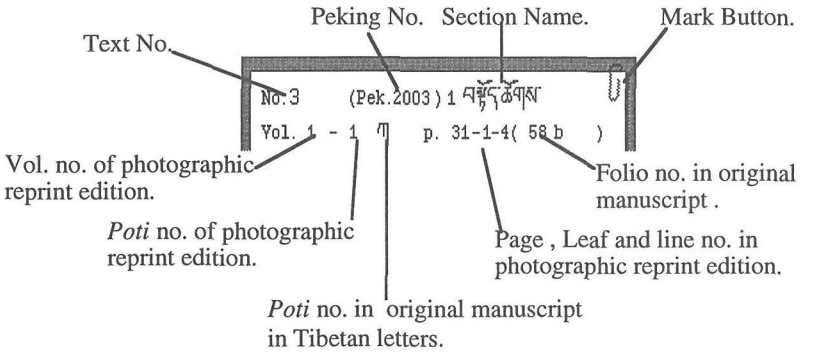
By clicking on the name of each section, the first card containing text appears.

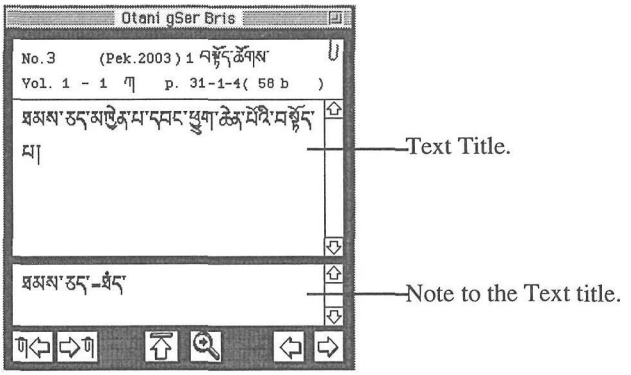
Cards Containing Text

Data for each text is contained within one card. Each card contains the following data.

Text title
Note to the text title
Text number
Peking number
Volume number of photographic reprint edition
Poti number in photographic reprint edition
Poti number in original manuscript in Tibetan letters.
Section name
Page, leaf and line number in photographic reprint edition
Folio number in original manuscript

The following is the card for GMTG no. 3.





Conclusion

The creation of electronic catalogues of Tibetan language materials has only just begun and many problems still remain to be solved. There still exists in both the Otani Tibetan Collection Stack and the Otani gSer Bris Stack many parts which remain incomplete. In particular, there is pressing need to create a true sort function for the stacks. There will undoubtedly appear many other problems as the work of creating electronic Tibetan databases progresses. But it is our hope that our work of creating such databases will help advance the study of Tibetan culture.

* Members of the Project include: Ichijo OGAWA (chief), Michio KATANO, Nobuchiyo ODANI, Tsultrim Kalsang KHANGKAR, Kazuo HYODO, Yoshiro IMAEDA, Yoichi FUKUDA.

Further assistance was provided by: Hideki KATO, Akimasa YOSHIDA, Tetsuo HIROHAMA, Tetsuya YAMADA, Toru KITO, Yukinari MAKITA, Shigeki NAGAO, Tomohiro SAKURAI, Shin'ichiro MIYAKE.